DOCUMENT RESUME

ED 169 103 TM 008 513

AUTHOR Hall, Melvin
TITLE Issues in Higher Education Program Evaluation.
PUB DATE Mar 78
NOTE 25p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Case Studies; *College Programs; Curriculum Evaluation; *Evaluation Methods; *Evaluators; Group Unity; *Higher Education; Information Utilization; Institutional Characteristics; *Program Evaluation; Surveys; *Systems Approach
IDENTIFIERS *Evaluation Problems

ABSTRACT

A comparative case study of seven universities initially focused on: (1) whether and how program evaluation systems differ from other systems approaches, such as management information and program budgeting; (2) desirable degree of overlap between evaluative recommendations and administrative action; and (3) limitations of program evaluation. Site visits raised several issue statements. First, what was the effect of staffing--specifically, the use of outside consultants, and the method of selecting evaluators? Second, did institutional tradition, campus governance, and state coordination alter the evaluation? Third, were the effects of formal evaluation structures, (checklists, protocols, schedules, task forces) a healthy boost to either institutions or to subsequent evaluations? Fourth, was there an advantage to linking program evaluation systems to budget systems? An assessment of these issues led to the conclusion that formal systems in higher education were new and significant modifications of more traditional program review activities such as accreditation, national ratings, institutional research, and institutional self-study. (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Melvin Hall

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ISSUES IN HIGHER EDUCATION PROGRAM EVALUATION

MELVIN HALL

SANGAMON STATE UNIVERSITY

ED169103

TM008 513

Preface

Systematic evaluation efforts developed especially for campus level program evaluation have become a common feature of many universities. In general, these evaluation systems have sought to operationalize some notion of accountability, consumer protection, or responsible planning. Evaluation had been a standard feature of campuses for several years, these newer structures were merely more visible systems of providing information for and faculty participation in administrative decision making. This new thrust also sought to combine evaluation with existing budgetary, personnel, and managerial systems.

The current discussion is grounded in a survey of several previously established review activities found in higher education. These more traditional review activities included institutional and program accreditation, national ratings, institutional research, operations research, institutional traditions, and institutional self-study. These traditional activities have been contrasted with systematic efforts at several university campuses across the United States. Campuses included in the study are:

University of Washington - Seattle
University of Oregon - Eugene
University of California - Berkeley
University of Michigan
University of Minnesota
State University of New York at Albany, Buffalo
University of Illinois - Urbana

Site visits to these campuses led to the development of several issue statements. The current paper is organized around four of these issues, including:

"Printed in U.S.A."

1. Was the staffing of evaluation systems as important a consideration as it appeared? Did it matter if the evaluation team was completely in-house? Was the level of involvement of faculty and student populations the crucial determinant in staffing? Were there important implications for evaluation models where staffing was prescribed?

2. The elements of institutional traditions or mystiques, campus governance structures (both internal and external), and state coordination combined to influence evaluation systems, the role of evaluation, and associated sanctioning processes. Did these elements alter the resultant evaluation?

3. Were the effects of elements of formal evaluation structures (checklists, protocols, schedules, task forces, etc.), a healthy boost either to institutions or subsequent evaluations?

4. Were there identifiable advantages to joining formal evaluation systems to budget systems in order to form a planning system? Did this structure have implications for the stature of operational or short term planning?

ISSUES IN HIGHER EDUCATION PROGRAM EVALUATION

Introduction

Assessment of the desirability, feasibility, and effectiveness of campus evaluation systems depends in part upon identifying whether these attempts at program review were a real and significant new evaluation force or merely a rearrangement of old activities. This review supports the notion that formal systems of program evaluation in higher education were new and significant modifications of more traditional program review activities.

This is not to imply that program reviews conducted in this formal mode were not in large part made up of procedures used previously in traditional evaluation activities. One need only examine two or three evaluation systems to discover that the differences in systemic reviews were not the result of dramatically new techniques or novel approaches. These formal systems do, however, reflect several departures from traditional evaluation efforts.

Among the more obvious departures, systematic efforts were broader based, cut across disciplinary and departmental borders, involved multiple rather than single or dual sanctions, followed contemporary notions of popular values, and recast the roles of many central campus figures. The writings which marked the establishment of these systems helped to make clear the altered perceptions of evaluations. These perceptions were demonstrated by statements of what the "system" might accomplish.

Goals of the New Systems

At the University of Michigan program review was expected to provide a broader base of information than was previously available to administrators in making decisions. The Program Evaluation Committee Report of October 1973 indicated that:

> At present, budgetary constraints from both state and federal sources force deans, department heads, directors and executive officers to make hard decisions regarding present and future operations of the University. These hard choices are made today based upon a variety of information and decision processes. One objective of the Program Evaluation Committee is to identify a set of procedures which may be of assistance to the key decision-makers as they focus upon an issue for allocation of limited fiscal resources. (p. 4)

A further aspiration of the University of Michigan's system of evaluation was imbedded in the following statement. "The question is not whether the University should conduct evaluative activities; rather, the question becomes whether the University should undertake evaluations the way that they are presently conducted or in a more efficacious manner." (University of Michigan, p. 4)

Other goals for systematic review efforts were enumerated by the University of Minnesota. These included:

1. To insure that each program conducts a periodic self-assessment of its status.

2. To allow each department the opportunity to present in terms of data and descriptive material its own perspective on its role in the University.

3. To provide the deans of the Graduate School and the Institute of Technology a data base and supplementary information to aid them in internal planning, in preparation of budget requests, and in meeting the requirements of accountability. (Brodbeck, 1974)

In addition, the systematic reviews of the University of Minnesota attempted to consolidate the University's approach to academic planning; to facilitate and sharpen periodic reviews; and to assist the faculty in examining and evaluating their program fields aided by university colleagues in related fields and by external visitors from other institutions and pertinent enterprises (Graduate Program Review Statement, revised September, 1973).

Many of these goals of formal program review could have been accomplished by less formal evaluative activities. The first goal stated by the University of Minnesota perhaps best summarized the perceived advantage of the newer "formal" structures over "traditional evaluation." Ratified formal procedures provided an increased measure of assurance that evaluative activities would be carried out regularly and in a prescribed manner.

Many program evaluation systems made use of administrators, program faculty, students, faculty from outside the program area, and often consultants external to the university. With this broad participation more perspectives were expressed, although discussions were more diffuse.

Relation of New Systems to Other "Systems" Approaches

These systems of review contrasted sharply with other "systems" approaches applied to higher education, such as management information systems and program budgeting systems. In program review systems, communication patterns were complex and intertwined. There was a sense in which these patterns recreated the patterns of communication typically found in the university's normal routine. Other "systems" approaches sought to alter and streamline these communications patterns making them regularized, coded and efficient. Program review allowed respondents to step out of traditional communication patterns with the expressed intent of describing the status quo, not replacing it. Replacement came later as a judgment activity.

A capstone review activity used at the University of Michigan perhaps best illustrates this point. At the University of Michigan the Vice-President for Academic Affairs participated in a capstone briefing near the end of program review activities. This briefing also involved the college dean, department chairpersons, and staff from the office of budgets and planning. This briefing was an opportunity for the Vice-President for Academic

Affairs to become familiar with college needs and future directions.

In a sense, this briefing was different from normal communication patterns because the chain of command called for discussions between department heads and deans, between deans and the vice-president, but not for group discussions between these parties. This variation did not seek to replace either the established communication patterns or the chain of command however; it merely accentuated these patterns, reduced the slippage and concentrated discussion on self-study results.

Limitations of the New Systems

Numerous limitations of the systems approach to program evaluation have also become evident. Program review systems tended to emphasize participation of faculty, students, administrators, and community groups. As a result committees or task groups were the dominant format. The committee approach often did not pinpoint administrative deficiencies. Committee members, through negotiation and compromise tended to generalize administrative deficiencies to other aspects of program functioning. Many assumed a less personalized systems approach which concentrated on data inputs and outputs would have provided more definitive and prescriptive recommendations to overcome administrative deficiencies.

Secondly, program review systems, again because of their dependence on committee work, reflected uneven effort within the range of topics covered in the review. The dynamics of committee functioning accounted for some of this unevenness, but it is difficult to separate this from the fact that all committee members do not function with the same level of proficiency.

A third major limitation of program review systems was that major effects of program review activities were not observable or, even when observed, reportable. Many individual faculty members had only meager information about

the efforts of the other faculty in their departments. Much of the effect of program review has been the result of the familiarization of faculty with program activities of which they had formally had little knowledge. In the University of Michigan example, a major result of program review came from the department head being involved in the discussion between the dean and Vice-President for Academic Affairs or the Vice-President for Academic Affairs observing interpersonal dynamics, between deans and department heads. Such small deviations from normal operating procedures are difficult to spot and to measure or describe, but may significantly alter the units evaluated.

A fourth limitation observed nationally has been the failure of these evaluation systems to meet the expectations held by participants. The wide scope of these systems made such varied demands on description and analysis activities that it was impossible to cover all the values held by participants. Consensus did not permit all view points to be appropriately portrayed, thus several points of view were not reflected in final evaluation reports.

A rifth limitation was illustrated by the failure of evaluation committee attempts to propose solutions incompatible with normal funding patterns. Because evaluation systems were based on the development of interdisciplinary solutions, there was considerable tension between global images developed by committees, and specific recommendations which were drafted.

Many questioned whether such systems were controlled by faculty, administrators, or both. Bennis (1976) and others illustrated, that the forces acting on higher education tended to fragment the authority of administrators, leaving them less effective and more vulnerable. Administrators sought to obtain support without further erosion of their control of university activities. As a result, administrators tended to use program review systems as a way to broaden their base of input and win support for administrative decisions.

The wider involvement which came with program review systems caused increased public scrutiny of administrative decisions. In addition the struggle for control in higher education led to demands that evaluation recommendations be reviewed by both faculty and administrators before adoption or widespread circulation. The broad involvement, which characterized these review systems, created multiple roles for individuals involved in program review. Faculty members became linked to the administration, coordinating board or whatever agency sanctioned the review process. Thus as program evaluation systems rewarded compliance with campus values, the faculty, student or community members involved in evaluation became messengers of those values. An institution clear in its mission and goals was certain to have exemplary models as participants in the program review process, and through recommendations and analysis of the committee campus values were perpetuated.

Finally, systems of program evaluation aggravated the overlap of administration and faculty because of the involvement these systems fostered. Having spent considerable time and energy in committee work drafting evaluative comments and recommendations, faculty typically became desirous that evaluation material be used, not misused. One outcome of this desire was a struggle for control over implementation of recommendations. Ego involvement often dictated that a faculty committee move into the political/administrative arena to support, defend, and clarify its comments, documentation, and recommendations, even though the committee had known all along that its status was strictly to be advisory.

Issue Statement No. 1

> Was the staffing of evaluation systems an important consideration? Did it matter if the evaluation team was completely in-house? Was the level of involvement of faculty and student populations the crucial determinant in staffing?

Importance of Staffing. The staffing of evaluation systems defined patterns of communication and participation, human resources available, and perspectives involved in the reviews. It is not clear whether this influence was exclusively the result of the selection criteria which guided the staffing of program review committees, or one of several other factors which influenced program evaluations. Not the least of these factors was the standard administrative structures which typically organized central administrations. Because the chief academic officer was also usually the chief budget officer, the location of evaluation staff within the university hierarchy was critical. In addition the perceived distance of evaluation from the chief academic officer determined the credibility and perceived autonomy of the review system. Because evaluation typically begins with academic programs, more specifically the expensive graduate degree programs, additional constraints are imposed on the program review activity by oragnizational dynamics.

Beyond situational factors, other stable influences mediated the staffing question. Dressel (1976) has illustrated the inappropriateness of using standing committees in institutional self-study or program review. Standing committees tended to be inhibited and ego involved with maintaining the status quo. While this inhibition did not make rational review impossible, it nonetheless substantially reduced committee effectiveness. Standing committees spent valuable time covering old ground and too often relied on old channels for soliciting information. These committee were often unsuitable for the tasks of institutional self-study and program review. Standing committees because of prior involvements had a jump on the identification of issues and information sources, but this was often negated by reduced objectivity, reduced creativity and the binding effect of strong internal interpersonal relationships and roles.

Executive and policy committees also were shown to be inappropriate for self-study activities. It should be noted, however, that often the appointment of a new coordinating committee for institutional review activities posed a threat to the established committees and their traditional or statuatory responsibilities. It was often the case that systematic review activities all but eliminated the role of these committees. In several cases a special central office or committee was used to coordinate all evaluative activities associated with periodic or systematic program review.

This central coordinating committee was typically an internal committee, coordinating design, implementation, and analysis of evaluation activities. Such activities included the efforts of subcommittees, task forces, administrators and institutional staff. Below this gross level, models for systematic review varied with program constituents and other forces outside the program or institution. The pattern of involvement was determined by the type of review involved, whether it is primarily for budget reallocation, long range planning, or accreditation. However, there were several general parameters which shaped the staffing issue.

Use of Outside Consultants. A focus of considerable attention was on the value of outside consultants in the program review process. The issues regulating the use of external consultants were two: first, the evaluative questions to be asked and, second, the available resources. Models ranged from a modified accreditation approach which used external consultants and site visits, to a wholly internal approach which used no regular involvement of outside consultants.

A further and overlapping consideration was the audience of the evaluation and its requirements. The audience dictated which evaluative perspective was most relevant. State coordination was more likely to involve com-

parison with other institutions both in state and out-of-state because it required assurances that public needs be efficiently met. Exclusively internal reviews had to grapple with questions of internal allocations and hence had to be concerned with how central a program was to the university's scope and mission regardless of its state or national reputation.

Cost was the second major factor affecting the use of outside consultants. For example, the State University of New York at Albany used external panels extensively for review of graduate programs. Consultants were primarily drawn from the eastern seaboard and typically spent one to two days on campus. Each of these panels of three scholars external to the university represented an average university investment of $1500. In an institution committed to reviewing 10 to 20 percent of its programs annually, external consultants may cost from 10 to 40 thousand dollars annually.

Often institutions attempted to combine internal and external reviewers. Various models of this sort existed. These ranged from appointing single external consultants to work with internal committees, to choosing independent internal and external committees coordinated by a third internal committee.

Selection Strategies. Perhaps no other element of formal evaluation structures more accurately reflected the prevailing style of an institution than the method of selection of evaluation participants. Models of selection ranged from constituency elections to the appointment of "faculty stars" by chief academic officers. Equally reflective of institutional style was the question of involvement and selection of students and community participants. Precedents existed of near-student dominated committees—for example, The Study of Graduate Education at Stanford—as well as of systems involving no formal student input. The involvement of faculty, students, administrators,

and others in program review has tended to follow the established institutional pattern of involvement for these groups.

An Example from the University of Illinois. The system of program review at the University of Illinois - Urbana was named the Council on Program Evaluation. (COPE) The operations of the COPE Council showed clearly several effects of staffing patterns on program evaluation systems. With regards to staffing, COPE was a two-tiered system. The coordinating tier included an appointed faculty committee and administrative staff support. This committee was appointed by the Vice-Chancellor for Academic Affairs who was the chief academic officer of the campus. Appointments to the council were made with the concurrance of the governance system, and included both faculty and student members.

The second tier of the evaluative system operated at the department or unit level and worked as either a departmental self-study committee or a COPE appointed task group composed of university faculty and students from both inside and outside the department. No external consultants were routinely used in program review.

Thus the COPE system was characterized by the automatic involvement of the Vice-Chancellor of Academic Affairs, an independently selected coordinating committee, student representation, single rather than joint reviews, and infrequent use of external consultants.

These parameters of staffing illustrated and defined a system of power relationships. The Vice-Chancellor having been a driving force in the establishment and operation of COPE, maintained appointment powers as COPE was implemented. These powers were mediated only by the concurrance which he was to seek from governance. The establishment of the chief academic and budget officer as a central figure in the selection process increased the planfulness

of COPE selection and operation.

Planfulness and faculty involvement in large part may be credited with contributing a broad planning and educational base to the COPE activity. Faculty participation fostered by COPE increased campus realization of the far-reaching effects of evaluation and management decisions. The ultimate goal of replacing year-to-year expediency with cohesive analysis, planning, and management was forwarded by this staffing arrangement.

A domino effect was also observable in COPE's staffing pattern. An obvious side effect of the central role of the Vice-Chancellor of Academic Affairs in establishing and maintaining COPE, was the undiluted influence which he had over the features of the review structure. As stated earlier, these influences had to co-exist with institutional traditions and the prevailing political climate.

Another factor in the determination of the pattern of participation was the interlocking effect of COPE affiliation. COPE membership required status as a respected member of the campus community. Senior campus citizens whose high professional standing was commonly accepted were nominated to COPE. The cohesiveness of campus perspectives were reflected in these nominations. At Urbana, the campus "sees itself as the pre-eminent publicly-supported graduate-level institution within the state . . . (and further) sees itself as distinct from other state-supported institutions because of its emphasis on research and seminal scholarship."

The disciplinary and rank cross-pollination which did exist supported an appearance of balance essential for the operation of a central campus evaluation activity. As the most obvious element of the evaluation process COPE was the most likely focus of attacks on evaluative recommendations. COPE had to appear to be balanced with respect to disciplines and perspectives

found within the institution.

While a cohesive committee perspective on preferred academic style facilitated evaluative discussions and multiplied the reliability of evaluative judgments, a point of some concern was that other perspectives became disenfranchised. It was important to maintain a threshold of agreement on campus missions, a threshold which was neither narrow, or totalitarian. The common perception on the Urbana campus was that COPE was not unduly restrictive. But some have speculated that COPE reflected a rigid fragmentation of academic pursuits on the campus. COPE has been accused of supporting an implicit junior partner relationship for many departments, based on the established scholarly values of the campus. The crucial question, of how much cohesion was necessary for productive and reliable evaluations was never completely answered although it was clear that campuses which had never clarified their mission were less receptive sites for systematic program review.

Two characteristics of COPE mediated the effects of cohesive campus values. COPE occupied a fairly nebulus zone between faculty and administration. As a result it was unclear what status should be ascribed to COPE. Secondly, the very nature of institutional self-study usually prevented COPE from uncovering unknown weaknesses in the campus armor. As a result COPE tended to influence the tenor rather than the substance of campus debates.

Faculty participation in program review increased participation of the faculty in university affairs and stimulated the sharing of information. The primary success of faculty participation was the broadening of investigation, dialogue, and human resources. Such participation did considerably less towards enfranchising faculty on major institutional decisions. Many in fact argued faculty involvement should never erode administrative prerogative. And in fact, COPE did not result in a totally participatory system of decision making.

COPE represented faculty participation in information gathering, problem definition, and even alternatives identification and analysis, but not in allocation, implementation, or follow-up.

Because of its peculiar stance COPE was also vulnerable to co-option and thus to reduction of the benefits of broad participation. In COPE co-option resulted from the unquestioning harmony which linked COPE to the administration which was to implement or ignore it's evaluative recommendations. Co-option was reflected in the degree to which the evaluative group speculated, not on the wisdom but marketability of its recommendations. The committee did not want to spend hours deliberating recommendations which would be buried in stacks of reports housed in attic storage. Thus, the temptation was great to anticipate the leanings of campus administrators. COPE tried to reduce this possibility by recruiting strong campus figures who collectively had stature independent of recommendation implementation. Yet, because committee stature was enhanced by the serious attention given to its recommendations, the danger of co-option persisted.

Co-option posed another threat to the quality of institutional self-evaluation since the redistribution of knowledge and general stimulation which should accompany self-study was easily choked. The expanded consciousness of departmental faculty expected from self-study activities needed the support and definition of public statements. If the coordinating committee became too closely aligned with administrators it rendered itself incapable of giving this support. The staffing patterns of formal program review systems therefore had much to do with whether the subsequent review was geared for narrow administrative audiences, and whether reviews encouraged fragmented perspectives, repeated traditional campus values or promoted academic growth.

Issue No. 2

> The elements of institutional traditions or mystiques, campus governance structures (both internal and external), and state coordination combined to influence evaluation systems, the role of evaluation, and associated sanctioning processes. Should these elements have altered the evaluations?

Traditions, mystiques, and other features of the higher education milieu encouraged campus cohesion. As discussed earlier, this cohesion was a necessary though not sufficient condition for viable program evaluation. Specifically campus traditions of student involvement, faculty participation and administrative prerogative were central staffing parameters.

Weaknesses of Formal Evaluation Programs. One area in which campus cohesion weakened formal evaluation activities was in the establishment of the evaluative criteria to be applied in the reviews. In establishing the criteria by which programs were to be assessed, it was important that long standing campus mores be put into perspective. Restriction of the range of criteria and the subtitution of campus traditions for evaluation criteria together imposed rigid controls on program reviews. Accepting quality research as a positive criterion, and further restricting quality research to scholarly academic research clearly reduced evaluative activities to very narrow activities based on a predesigned system of valuings where only traditional esoteric research was to be highly regarded.

In evaluation, and particularly in institutional self-study, it was very easy to allow, ease of measurement, strength of tradition, or influence of governing boards to determine not only what constitued evidence but also how that evidence was to be valued. This was true even though popular notions of accountability assumed very cohesive institutional goals which could be "behaviorably" defined.

This perspective on the relationship between campus cohesion and

evaluation supports several observations regarding those formal evaluation structures which have tended to promote traditional perspectives. The first such observation is that such systems tended to be internally oriented. Given the two primary evaluation questions of internal and external comparison, traditional structures were more likely to emphasize internal questions. Rationales for controversial recommendations were steeped in the rhetoric of traditional campus values and surrounded by an air of matter-of-factness. Often this independence was manifested in attempts to make the evaluation activities a regular cyclical campus activity. This is not to say that all cyclical evaluations were tradition dominated, but merely that tradition dominated systems sought the stability and security which regularization brought.

Beyond the campus cohesion question, several additional elements of campus expression connected to institutional self-image and mystiques or traditions are noteworthy. As the flagship campus of the University of Illinois, the Urbana campus had been considered by many to be the hub of state supported higher education. Traditions of scholarly research, supported an image of campus integrity. One outcome was the progressive stance of campus administrators which found expression in program evaluation. Urbana administrators, operating on a do-it-to-yourself-before-it's-done-to-you philosophy, began program evaluation without particular external pressure for program review. In establishing COPE, these administrators were able to further reassert the integrity of the campus community. By being among the first in the state to engage in systematic self-evaluation Urbana assured its independence; COPE would shape future state initiatives in program evaluation rather than being shaped by them. This has in fact occurred. The Illinois Board of Higher Education's state wide program evaluation effort has required no significant changes in COPE.

Urbana's second independent motivation for program review was the

campus administration's incrementalist perspective on change. Following campus precedents, COPE was created as a tool for reallocation, and was not meshed with either the budget cycle or other budget mechanisms. Its impact was intended to be steady and reasoned. The subtle pressure of peer review was to be its major power. COPE was intended to provide the Vice-Chancellor for Academic Affairs the information he required for administrating the campus, but perhaps more important there was hope that through the processes of faculty input enough information could be shared within departments to make administrative action less frequently required. One central campus administrator complained that administrators were forced to give 90 percent of their time to what amounted to 10 percent of their responsibility. Systems like COPE had the advantage of studying their subject without critical urgency. The overall goal was to facilitate incremental movement toward change and avoidance of the impass, and entrophy found in organization unprepared to deal with unforseen circumstances.

Issue Statement No. 3

> Were the effects of elements of formal evaluation stuctures (checklists, protocols, schedules, task forces, etc.), a healthy boost either to institutions or to subsequent evaluation?

I do not share the opinion that formal evaluation activities in higher education were cosmetic changes designed to insure commitment to program review.

Although formal evaluative structures used fairly traditional instruments and strategies they added a longitudinal dimension. Typically program reviews were repeated every five to ten years, within every part of the university. This certainty has caused institutions to look for ways of using the vast amounts of comparative data available. Programs have been asked to identify peers with whom they might be compared, and program recommendations

made five to ten years earlier became starting points for a second round of evaluations.

Anticipation of evaluation carried additional momentum and influence when it included knowledge of the format to be used. The organization of review formats provided a model for standing committees, and filing systems. However, many expressed concern that program functions would become too standardized. Many felt too much emphasis was on how to get the program into the compartments defined by the standard format.

For campuses with a fairly cohesive notion of what they are about, these protocols and questionnaires further solidified an already pervasive self-image. Programs were bombarded with subtle pressures to conform to the standard image implied in the formats, and to the criterion they reflected. Sometimes altering the delicate balance of perspectives caused shifts in programs as they attempted to become aligned with the perceived central mandate.

At Urbana, for example, the establishment of COPE ushered into existence a new survey of student majors (Program Evaluation Survey—PES), a task force listing of categories of evaluation, a 14 category departmental self-study survey, standardized faculty activity summaries, seven standard tables for depicting instructional costs figures and faculty effort statistics, a department chairperson evaluation system, a standard unit resource outline, and numerous special data analysis formats. Each of these elements carried a coded image of an ideal department. Faculty activity summaries contained seven major headings, including such measures of national visability as editorial duties, regional/national offices held, major honor and professional society memberships, scholarly output, a recent research support summary, years of experience, publications, and a university service committee summary.

These protocols operationalized dominant campus values and rein-

forced these values because, once released, they tended to be only selectively revised in response to pointed criticisms. Thus over several years departments and department faculty came to expect that these seven categories of faculty activity would be important in COPE deliberations.

COPE has been criticized for allowing its standardized programatic inputs to result in a depersonalization of the review process. Department chairpersons have judged COPE less responsive than might be desired. Dissatisfaction was higher in departments which viewed themselves as not fitting the standard campus mode. Attempts have been made to alter this image apparently with some success judging by comments from campus governance, indicating confidence in the ability of COPE council members to weigh carefully the implications/limitations of such standardized information requests. In addition some attempt was made to supplement formal responses with direct departmental information gathered by both COPE staff and council members.

Although the criticism of inflexibility has been addressed, two other questions remain unanswered. First, is the question of whether standardized formats allow units to portray the actual flow of their transactions, or only stimulate the generation of bureaucratic dialogue. On this issue two dominant perspectives have been articulated, both in support of standarized formats. One position was premised on the notion that units had a story to tell, which questionnaires and protocols simply illicit. Program statements, then were not unduly influenced by questionnaires. A second opinion contended that quite the contrary occurred. In this perspective programs were viewed as having interacted with protocols or questionnaires to develop a statement. The implication was that although instrumentation may affect the statements illicited, standardization allowed comparison.

Impact of Systems. Systematic reviews have not uncovered new problems for campus consideration. What such systems have tended to do is twofold. First, the systematic collection of data has both allowed clearer statement of submerged campus issues and supported enlightened statements regarding the pervasiveness of these issues either within a single unit or through several departments. Secondly, by getting these problems areas into the public arena, program evaluation has facilitated administrative action. In some cases this stimulation has forced administrations to deal with smoldering issues.

While program evaluation stimulated analysis and dialogue in one case, it has limited communication in others. Limitation occurred when standard protocols reduced the need for contact between evaluative units and the departments being evaluated. Institutionalized formats which came to be viewed as refined instruments for data solicitation tended to replace the hands-on stimulation of some traditional evaluation activities. They also reduced conversations to explanation of the various forms and procedures involved.

Related to the reduction of personal communication is a problem of insufficient communication, which was also linked to standardized data formats. In an attempt to broaden the applicability of protocols, the specific items used tended to be less focused than might be expected. Hence useful and unuseful information was received laced together. Standardized questionnaires worked as lightening rods and attracted some troubling, some unuseful, and some jeopardizing information in their attempt to be complete. Distilling these reports resulted in a time lag for circulation of recommendations.

Another element of evaluation influenced by formal structures was confidentiality. Collection of standard data elements on every unit created

a ready market for unofficial release of program data which lead to public comparisons based on narrow slices of evaluative data. In the case of COPE, however, news leaks had been less a problem in the more standardized self-study format than in the earlier task force system, perhaps because the standardized self-study data was more routinely releasable.

One unanswered question was the ability of these newer evaluation systems to reconcile what was generally known about program quality with what was learned about the philosophy underpinning departmental activities. It was not clear from available evidence if common formats were capable of generating the information required for making this synthesis. It was unclear how much of the synthesis was the result of the combined insight of COPE members and was hence dependent on quality COPE membership.

Issue Statement No. 4

> Was there an identifiable advantage to joining formal evaluation systems to budget systems in order to form a planning system? Did this structure alter the stature of operational or short term planning?

A primary benefit of linking program evaluation systems with planning or budgeting systems was that the evaluative activity was thereby reinforced. The power to reward departments with a higher priority for new monies or increased physical space added much stature to program review activities, and created an advantage for systematic evaluation activities over earlier evaluation efforts which had been forced to operate on the negative sanctions associated with peer pressure. The linking of evaluative and other functional systems imparted positive rewards to balance these otherwise negative sanctions. This linking also increased faculty involvement in important decisions, though this still largely depended on the direction provided by evaluation participants.

In addition, the linking of evaluation with other systematic efforts set dominant campus perspectives in an even stronger position. The imposition of powerful perspectives was hard to avoid when the criteria were used in program review, long range planning, and budget decisions. Integretated systems of this type had great potential for tightening up campus planning and had to be closely monitored because of their repressive potential.

In summary, linking evaluation to other activities has provided the potential for preventative maintanence. With systematic evaluation problem areas have been addressed incrementally by various segments of the university before they had the opportunity to develop into university-wide concerns. The primary difficulty has been that program evaluation got further involved in academic management. On campuses which moved too far in this direction, evaluation systems developed which attempted to replace the statutorily created mechanisms for university management.

SELECTED BIBLIOGRAPHY

Unpublished Memorandums, Manuscripts, and University Documents.

Brodbeck, May. University of Minnesota Graduate School Memorandum, February 1974.

Scriven, Michael. Program Evaluation at U. C. Berkeley, Unpublished Manuscript, July 1975.

State University of New York at Albany Office of Graduate Studies, Procedure for Review of Established Graduate Programs, March 15, 1973.

University of Illinois at Urbana-Champaign, Preliminary Report of the Study Committee on Program Evaluation (SCOPE), March 1972.

University of Illinois at Urbana-Champaign Faculty Senate, Report of the Evaluation of the Council on Program Evaluation (COPE) by the Committee on General University Policy, May 4, 1977.

University of Illinois at Urbana-Champaign Office of the Vice-Chancellor for Academic Affairs, 1972-73 Academic Affairs Directive No. 6, July 19,1972.

University of Michigan Office of Budgets and Planning, Program Evaluation Committee Report, October 1973.

University of Minnesota, Department/Program Review Procedures Joint College of Liberal Arts - Graduate School, Protocol, June 1974.

University of Washington, Graduate School Memorandum No. 6 Draft VII, July 15, 1972.

Formal Publications

Bennis, Warren. The Unconcious Conspiracy, Amacom, 1976.

Carlsmith, Merrill J. et al. The Study of Graduate Education at Stanford, Stanford University, June 1972.

Hodgkinson, Harold and Myron B. Bloy. Identity Crisis in Higher Education, Jossey-Bass Inc.: San Francisco, 1971.

Hoenack, Stephen A. and Alfred L. Norman. Incentives and Resource Allocation In Universities, Journal of Higher Education XLV, No. 1, January 1974.

Dressel, Paul L. Handbook of Academic Evaluation, Jossey-Bass Publishers, 1976.